RAW SEQUENCE LISTING   DATE: 08/15/2002
PATENT APPLICATION: US/09/595,947C   TIME: 14:27:17

Input Set : **A:\Pto.raw**
Output Set: **N:\CRF4\08152002\I595947C.raw**

ENTERED

```
  1 <110> APPLICANT: ICARD-LIEPKALNS, Christine
  2        MALLET, Jacques
  3        RAVASSARD, Philippe
  4 <120> TITLE OF INVENTION: POLYPEPTIDES OF THE "BASIC-HELIX-LOOP-HELIX" bHLH
  5        FAMILY,
  6        CORRESPONDING NUCLEIC ACID SEQUENCES
  7 <130> FILE REFERENCE: ST96042AUS
C--> 8 <140> CURRENT APPLICATION NUMBER: US/09/595,947C
  9 <141> CURRENT FILING DATE: 2000-06-16
 10 <150> PRIOR APPLICATION NUMBER: FR96/15651
 11 <151> PRIOR FILING DATE: 1996-12-19
 12 <150> PRIOR APPLICATION NUMBER: PCT/FR97/02368
 13 <151> PRIOR FILING DATE: 1997-12-19
 14 <150> PRIOR APPLICATION NUMBER: US09/331,356
 15 <151> PRIOR FILING DATE: 1997-12-19
 16 <160> NUMBER OF SEQ ID NOS: 28
 17 <170> SOFTWARE: PatentIn version 3.0
 19 <210> SEQ ID NO: 1
 20 <211> LENGTH: 1460
 21 <212> TYPE: DNA
 22 <213> ORGANISM: Rattus norvegicus
 23 <400> SEQUENCE: 1
 24        gcaggtagcg agaggagcag tccctgggcc cccgttgctg attggcccgt ggcacaggca     60
 25        gcagcccggc aggcacgctc ctggtccggg cagagcagat aaagcgtgcc aggggacaca    120
 26        cgattagcag ctcagaagtc cctctgggtc tcaccactgc acagaggccg aggacccct     180
 27        ccgagcttct ttgctgcctc cagacgcaat ttactccagg cgaggggcgcc tgcagctcag    240
 28        caaaacttcg aagcgagcag aggggttcag ctatccaccg ctgcttgact ctgaccaccc    300
 29        gcagctctct gttcttttga gcccggagta actaggtaac atttaggaac ctccaaaggg    360
 30        tagaagaggg gagtgggtgg gcgtactcta gtcccgcgtg gagtgacctc taagtcagag    420
 31        actgtcacac cccccttcca ttttttccca acctcaggat ggcgcctcat cccttggatg    480
 32        cgcccaccat ccaagtgtcc caagagaccc agcaacctt tcccggagcc tcggaccacg    540
 33        aagtgctcag ttccaattcc accccaccta gccccactct cgtaccgagg gactgctccg    600
 34        aagcagaagc aggtgactgc cgagggacat cgaggaagct ccgtgcgcgg cgcggagggc    660
 35        gcaacaggcc caagagcgag ttggcactga gcaagcagcg acgaagccgg cgcaagaagg    720
 36        ccaacgaccg ggagcgcaac cgcatgcaca accttaactc cgcgctggat gcgctgcgcg    780
 37        gtgtcctgcc caccttcccg gatgacgcca aacttacaaa gatcgagacc ctgcgcttcg    840
 38        cccacaacta catttgggca ctgactcaga cgctgcgcat agcggaccac agcttctacg    900
 39        gccccgagcc ccctgtgccc tgtggggagc tgggaagccc gggaggggggc tccagcggcg    960
 40        actggggctc tatctactcc ccagtttccc aagctggtag cctgagcccc acagcctcat   1020
 41        tggaggagtt ccctggcctg caggtgccca gctccccatc ctgtctgctc ccgggcaccc   1080
 42        tggtgttctc agacttcttg tgaagggccc aaacaggccc tgggcggtgg gcgctggcag   1140
 43        aaagggaggg agtcagagct gtctgaaatg gaaggtagtg gaggcactcg agcatctcgc   1200
 44        cccttctggc tttcattagt caggtccctg atttaaccag gattcgcaca gttccttgct   1260
```

RAW SEQUENCE LISTING   DATE: 08/15/2002
PATENT APPLICATION: US/09/595,947C   TIME: 14:27:17

Input Set : A:\Pto.raw
Output Set: N:\CRF4\08152002\I595947C.raw

```
45        gctgtgcgtg cacaaaggac attgcaggct gatctcctct taaccctcct cagtgtggcc    1320
46        acctcaaact cccgctccaa gcagaggaga gccgtagcac taaatagttg ggagactccc    1380
47        atacttcctg gtgactccgc cctctttcaa atctgcgggc ctccaaccac cgctttctcc    1440
48        agagtgacct aatccagtgt                                                 1460
50 <210> SEQ ID NO: 2
51 <211> LENGTH: 24
52 <212> TYPE: PRT
53 <213> ORGANISM: Artificial Sequence
54 <220> FEATURE:
55 <223> OTHER INFORMATION: peptide fragment of bHLH protein
56 <400> SEQUENCE: 2
57        Ala Ala Thr Lys His Gly Met Gly Ile Gly Ala Gly Cys Gly Cys Ile
58        1               5                   10                  15
59        Asp Lys Cys Gly Cys Arg Tyr Gly
60                      20
62 <210> SEQ ID NO: 3
63 <211> LENGTH: 24
64 <212> TYPE: PRT
65 <213> ORGANISM: Artificial Sequence
66 <220> FEATURE:
67 <223> OTHER INFORMATION: peptide fragment of bHLH protein
68 <400> SEQUENCE: 3
69        Gly Gly Cys Ser Arg Asp Thr Tyr Thr Cys Ala Gly Gly Gly Thr Ser
70        1               5                   10                  15
71        Tyr Asx Gly Ala Tyr Cys Thr Thr
72                      20
74 <210> SEQ ID NO: 4
75 <211> LENGTH: 25
76 <212> TYPE: DNA
77 <213> ORGANISM: Artificial Sequence
78 <220> FEATURE:
79 <223> OTHER INFORMATION: primer
80 <400> SEQUENCE: 4
81        aaccttaact ccgcgctgga tgcgc                                           25
83 <210> SEQ ID NO: 5
84 <211> LENGTH: 18
85 <212> TYPE: DNA
86 <213> ORGANISM: Artificial Sequence
87 <220> FEATURE:
88 <223> OTHER INFORMATION: primer
89 <400> SEQUENCE: 5
90        cgcggtgtcc tgcccacc                                                   18
92 <210> SEQ ID NO: 6
93 <211> LENGTH: 6
94 <212> TYPE: DNA
95 <213> ORGANISM: Artificial Sequence
96 <220> FEATURE:
97 <223> OTHER INFORMATION: DNA sequence of E box
98 <400> SEQUENCE: 6
```
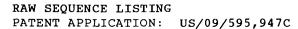
RAW SEQUENCE LISTING   DATE: 08/15/2002
PATENT APPLICATION: US/09/595,947C   TIME: 14:27:17

Input Set : A:\Pto.raw
Output Set: N:\CRF4\08152002\I595947C.raw

```
99        caggtg                                                              6
101 <210> SEQ ID NO: 7
102 <211> LENGTH: 6
103 <212> TYPE: DNA
104 <213> ORGANISM: Artificial Sequence
105 <220> FEATURE:
106 <223> OTHER INFORMATION: DNA sequence of mutated E box
107 <400> SEQUENCE: 7
108        tccgtg                                                              6
110 <210> SEQ ID NO: 8
111 <211> LENGTH: 214
112 <212> TYPE: PRT
113 <213> ORGANISM: Rattus norvegicus
114 <400> SEQUENCE: 8
115       Met Ala Pro His Pro Leu Asp Ala Pro Thr Ile Gln Val Ser Gln Glu
116       1               5                   10                  15
117       Thr Gln Gln Pro Phe Pro Gly Ala Ser Asp His Glu Val Leu Ser Ser
118                   20                  25                  30
119       Asn Ser Thr Pro Pro Ser Pro Thr Leu Val Pro Arg Asp Cys Ser Glu
120               35                  40                  45
121       Ala Glu Ala Gly Asp Cys Arg Gly Thr Ser Arg Lys Leu Arg Ala Arg
122           50                  55                  60
123       Arg Gly Gly Arg Asn Arg Pro Lys Ser Glu Leu Ala Leu Ser Lys Gln
124       65                  70                  75                  80
125       Arg Arg Ser Arg Arg Lys Lys Ala Asn Asp Arg Glu Arg Asn Arg Met
126                   85                  90                  95
127       His Asn Leu Asn Ser Ala Leu Asp Ala Leu Arg Gly Val Leu Pro Thr
128                   100                 105                 110
129       Phe Pro Asp Asp Ala Lys Leu Thr Lys Ile Glu Thr Leu Arg Phe Ala
130                   115                 120                 125
131       His Asn Tyr Ile Trp Ala Leu Thr Gln Thr Leu Arg Ile Ala Asp His
132           130                 135                 140
133       Ser Phe Tyr Gly Pro Glu Pro Pro Val Pro Cys Gly Glu Leu Gly Ser
134       145                 150                 155                 160
135       Pro Gly Gly Gly Ser Ser Gly Asp Trp Gly Ser Ile Tyr Ser Pro Val
136                   165                 170                 175
137       Ser Gln Ala Gly Ser Leu Ser Pro Thr Ala Ser Leu Glu Glu Phe Pro
138                   180                 185                 190
139       Gly Leu Gln Val Pro Ser Ser Pro Ser Cys Leu Leu Pro Gly Thr Leu
140                   195                 200                 205
141       Val Phe Ser Asp Phe Leu
142           210
144 <210> SEQ ID NO: 9
145 <211> LENGTH: 1330
146 <212> TYPE: DNA
147 <213> ORGANISM: Homo sapiens
148 <400> SEQUENCE: 9
149       cctcggaccc cattctctct tctttctcc tttggggctg gggcaactcc caggcggggg    60
150       cgcctgcagc tcagctgaac ttggcgacca gaagcccgct gagctcccca cggccctcgc   120
```

RAW SEQUENCE LISTING   DATE: 08/15/2002
PATENT APPLICATION: US/09/595,947C   TIME: 14:27:17

Input Set : A:\Pto.raw
Output Set: N:\CRF4\08152002\I595947C.raw

```
151         tgctcatcgc tctctattct tttgcgccgg tagaaaggta atatttggag gccttcgagg   180
152         gacgggcagg ggaaagaggg atcctctgac ccagcggggg ctgggaggat ggctgttttt   240
153         gttttttccc acctagcctc ggaatcgcgg actgcgccgt gacggactca aacttaccct   300
154         tccctctgac cccgccgtag gatgacgcct caaccctcgg gtgcgcccac tgtccaagtg   360
155         acccgtgaga cggagcggtc cttccccaga gcctcggaag acgaagtgac ctgccccacg   420
156         tccgccccgc ccagccccac tcgcacaccg gggaactgcg cagaggcgga agagggaggc   480
157         tgccgagggg ccccgaggaa gctccgggca cggcgcgggg gacgcagccg gcctaagagc   540
158         gagttggcac tgagcaagca gcgacggagt cggcgaaaga aggccaacga ccgcgagcgc   600
159         aatcgaatgc acgacctcaa ctcggcactg gacgccctgc gcggtgtcct gcccaccttc   660
160         ccagacgacg cgaagctcac caagatcgag acgctgcgct cgcccacaa ctacatctgg   720
161         gcgctgactc aaacgctgcg catagcggac cacagcttgt acgcgctgga gccgccggcg   780
162         ccgcactgcg gggagctggg cagcccaggc ggtcccccccg gggactgggg gtccctctac   840
163         tccccagtct cccaggctgg cagcctgagt cccgccgcgt cgctggagga gcgacccggg   900
164         ctgctggggg ccacctcttc cgcctgcttg agcccaggca gtctggcttt ctcagatttt   960
165         ctgtgaaagg acctgtctgt cgctgggctg tgggtgctaa gggtaaggga gagggaggga   1020
166         gccgggagcc gtagagggtg gccgacggcg gcggccctca aaagcacttg ttccttctgc   1080
167         ttctccctag ctgacccctg gccggcccag gcctccacgg gggcggtagg ctgggttcat   1140
168         tccccggccc tccgagccgc gccaacgcac gcaacccttg ctgctgcccg cgcgaagtgg   1200
169         gcattgcaaa gtgcgctcat tttaggcctc ctctctgcca ccaccccata atcccattca   1260
170         aagaatacta gaatggtagc actacccggc cggagccgcc caccgtcttg ggtcgcccta   1320
171         ccctcactca                                                          1330
```

173 <210> SEQ ID NO: 10
174 <211> LENGTH: 214
175 <212> TYPE: PRT
176 <213> ORGANISM: Homo sapiens
177 <400> SEQUENCE: 10
178   Met Thr Pro Gln Pro Ser Gly Ala Pro Thr Val Gln Val Thr Arg Glu
179   1   5   10   15
180   Thr Glu Arg Ser Phe Pro Arg Ala Ser Glu Asp Glu Val Thr Cys Pro
181   20   25   30
182   Thr Ser Ala Pro Pro Ser Pro Thr Arg Thr Pro Gly Asn Cys Ala Glu
183   35   40   45
184   Ala Glu Glu Gly Gly Cys Arg Gly Ala Pro Arg Lys Leu Arg Ala Arg
185   50   55   60
186   Arg Gly Gly Arg Ser Arg Pro Lys Ser Glu Leu Ala Leu Ser Lys Gln
187   65   70   75   80
188   Arg Arg Ser Arg Arg Lys Lys Ala Asn Asp Arg Glu Arg Asn Arg Met
189   85   90   95
190   His Asp Leu Asn Ser Ala Leu Asp Ala Leu Arg Gly Val Leu Pro Thr
191   100   105   110
192   Phe Pro Asp Asp Ala Lys Leu Thr Lys Ile Glu Thr Leu Arg Phe Ala
193   115   120   125
194   His Asn Tyr Ile Trp Ala Leu Thr Gln Thr Leu Arg Ile Ala Asp His
195   130   135   140
196   Ser Leu Tyr Ala Leu Glu Pro Pro Ala Pro His Cys Gly Glu Leu Gly
197   145   150   155   160
198   Ser Pro Gly Gly Pro Pro Gly Asp Trp Gly Ser Leu Tyr Ser Pro Val
199   165   170   175
200   Ser Gln Ala Gly Ser Leu Ser Pro Ala Ala Ser Leu Glu Glu Arg Pro

RAW SEQUENCE LISTING   DATE: 08/15/2002
PATENT APPLICATION: US/09/595,947C   TIME: 14:27:17

Input Set : A:\Pto.raw
Output Set: N:\CRF4\08152002\I595947C.raw

```
201                 180                185                190
202      Gly Leu Leu Gly Ala Thr Ser Ser Ala Cys Leu Ser Pro Gly Ser Leu
203             195                200                205
204      Ala Phe Ser Asp Phe Leu
205             210
207 <210> SEQ ID NO: 11
208 <211> LENGTH: 18
209 <212> TYPE: DNA
210 <213> ORGANISM: Artificial Sequence
211 <220> FEATURE:
212 <223> OTHER INFORMATION: primer
213 <400> SEQUENCE: 11
214      caacgaccgg cagcgcaa                                        18
216 <210> SEQ ID NO: 12
217 <211> LENGTH: 24
218 <212> TYPE: DNA
219 <213> ORGANISM: Artificial Sequence
220 <220> FEATURE:
221 <223> OTHER INFORMATION: primer
222 <400> SEQUENCE: 12
223      gcccagatgt agttgtgggc gaag                                 24
225 <210> SEQ ID NO: 13
226 <211> LENGTH: 60
227 <212> TYPE: DNA
228 <213> ORGANISM: Artificial Sequence
229 <220> FEATURE:
230 <223> OTHER INFORMATION: primer
231 <220> FEATURE:
232 <221> NAME/KEY: misc_feature
233 <223> OTHER INFORMATION: n=a or t or g or c
234 <400> SEQUENCE: 13
W--> 235 atcgttgaga ctcgtaccag cagagtcacg agagagacta cacggtactg gnnnnnnnnn    60
237 <210> SEQ ID NO: 14
238 <211> LENGTH: 20
239 <212> TYPE: DNA
240 <213> ORGANISM: Artificial Sequence
241 <220> FEATURE:
242 <223> OTHER INFORMATION: primer
243 <400> SEQUENCE: 14
244      agacgacgcg aagctcacca                                      20
246 <210> SEQ ID NO: 15
247 <211> LENGTH: 24
248 <212> TYPE: DNA
249 <213> ORGANISM: Artificial Sequence
250 <220> FEATURE:
251 <223> OTHER INFORMATION: primer
252 <400> SEQUENCE: 15
253      gctcaccaag atcgagacgc tgcg                                 24
255 <210> SEQ ID NO: 16
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 08/15/2002
PATENT APPLICATION: US/09/595,947C   TIME: 14:27:18

Input Set : A:\Pto.raw
Output Set: N:\CRF4\08152002\I595947C.raw

<u>Please Note:</u>
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:13; N Pos. 52,53,54,55,56,57,58,59,60